SECTION 11. SIMULATION AND CALIBRATION OF RIVERINE FATE AND TRANSPORT OF NUTRIENTS AND SEDIMENT

Contents

Figures

Tables

SECTION 11. SIMULATION AND CALIBRATION OF RIVERINE FATE AND TRANSPORT OF NUTRIENTS AND SEDIMENT

11.1 Introduction

Riverine water quality and physical parameters, including transport of nutrients and sediment in the watershed, are simulated using HSPF (Bicknell et al. 1997; 2001; Donigian et al. 1984; Johanson et al. 1980). The HTRCH module simulates heat and temperature; OXRX simulates oxygen and BOD dynamics; NUTRX simulates inorganic nitrogen and phosphorus; PLANK simulates phytoplankton and benthic algae; and SEDTRN simulates the entrainment, scour, and deposition of sediment, as discussed in Section 9.

Like the hydrology and land-based nutrient and sediment calibrations, all riverine calibrations except for temperature are rule-based optimizations where a parameter in the river simulation is linked to a specific calibration metric or set of metrics. These riverine calibration metrics were found heuristically through calibration experience, sensitivity tests, and trial and error. The calibration of temperature is a simple optimization, maximizing model efficiency with a single parameter.

Additionally, regional factors, which adjust input edge-of-stream loads, are applied in the calibration process to account for unsimulated fate and transport processes which result in differences between the simulated and observed nutrient loads in the river simulation.

11.2 Water Quality Simulation

Constituents of the nutrient calibration are discussed under the subheadings below. For more detailed information on HSPF calibration, see Bicknell et al. 2001. The simulation of sediment was described previously in Section 9.

11.2.1 Temperature Simulation

The temperature of four soil layers is simulated in HSPF. The soil and the water contained within it are assumed to be at the same temperature. Water temperature has significant effects on reaction rates for the nutrient simulation (see Section 10). The surface and upper soil layer temperatures are linear functions of the air temperature. The intercepts of these functions vary monthly. The lower soil layer temperature and groundwater temperature are specified as monthly inputs. The simulated riverine heat balance processes include inflow, outflow, precipitation, evaporation, shortwave radiation, longwave radiation, and sensible heat transfer from the bed.

Initially, a calibration method was developed that was similar to the hydrology calibration in that upstream land parameters were calibrated to downstream gages. Ultimately, this method produced calibrated variability in the soil temperatures that affected the nutrient simulation. It is unlikely that this variability was related to real differences in the landscape, so the method was discontinued after using it to inform static parameterization.

The calibration method theorized that a relationship should exist between the average air temperature of a region and the parameters used to describe the soil temperature. Initially, average groundwater temperature was set to the average temperature of each land-segment. The groundwater temperature was given a seasonal variation with a winter/summer cycle of 10 degrees Fahrenheit (°F). Initially, the parameters controlling the surface layer temperature were set for each land- segment and month such that when the air was at the monthly average, the soil was also at the monthly average in the surface layer. The upper layer was set similarly, but with a seasonal variability at 90 percent of the seasonal variability of the surface layer.

As a test of the parameterization, the initial values were then calibrated to match observations in the stream. As noted above, this produced undesirable variability in soil temperature but revealed useful patterns. The calibration increased spatial variability but did not significantly change the underlying relationships for the surface and upper layer. Therefore the parameterization of the surface and upper level was accepted. In the case of groundwater, the set of calibrated groundwater temperatures by land segment was unrelated to the average air temperature by land segment. In the final calibration the groundwater and lower soil layer temperature were set to 52 °F as an average for all land-segments which was also allowed to vary by 10 °F between winter and summer. With the soil temperature parameterization held constant, the river was calibrated for temperature as described below.

11.2.2 Other Water Quality Constituents

Constituents other than flow, heat, and fixed solids are simulated variously in the HSPF modules OXRX, NUTRX, and PLANK. Each water quality constituent has a primary module, but it can be modified in other modules. For example, biochemical oxygen demand (BOD) is primarily simulated in the oxygen subroutine group OXRX, but it is also modified by phytoplankton dynamics in PLANK. Therefore, these will all be discussed together. Generally, all constituents are subject to advection, which is modeled as a function of flow with no user-controlled parameters. Particulate constituents are also subject to deposition and loss to the system. Generally, deposition is a function of water depth and deposition velocities, which are parameterized separately for each particulate water quality constituent including refractory organic material, algae, and nutrients attached to inorganic particulates.

The dissolved oxygen simulation includes reaeration, saturation, BOD decay, benthic demand, algal production, and nitrification. BOD includes deposition, decay, benthic release, and phytoplankton death.

Simulated nitrogen consists of ammonia/ammonium, nitrate/nitrite, and organic forms. Organic nitrogen includes BOD, refractory organics, and plankton. Transformations between the forms include nitrification, sorption/desorption of ammonia, BOD decay, and phytoplankton growth and death. Changes in total mass include denitrification, volatilization of ammonia, settling of particulate forms, benthic release, and benthic scour. Phosphorus simulation is similar to nitrogen and shares BOD and phytoplankton-related masses, but there is no counterpart to the nonconservative simulation of denitrification or ammonia volatilization.

11.3 Calibration

An automated method of water quality calibration was developed for the Phase 5.3 Watershed Model. The water quality method is similar to the hydrology calibration and the land nutrient and sediment calibration in that simulation parameters are paired with calibration metrics such that each parameter can be optimized to a unique set of metrics. There are, however, several differences.

The river calibration has advantages over the other calibrations in that the river simulation does not require a rerun of the associated land-segments as the software distributes the land loads and applies BMPs. The function evaluation time, or model run time, is very much shorter. In addition, river parameters apply to river-segments, so there is no issue with parameters applying across multiple river systems as with the hydrology calibration. The ability to separate basins allows for greater parallelization of runs.

Nested stations were considered by assigning a relative weight to each downstream station for each segment. The weight function was equal to the number of observations above the limit of detection and discounted by 90 percent when downstream of another water quality station. The weights were then scaled to add to 1.0 for each river-segment. For example, suppose a river-segment has two downstream water quality stations. The upper station has 50 observations, of which 20 are below the detection limit. The lower station has 400 observations, of which 200 are below the detection limit. The upper station is assigned a weight of 30. The lower station is assigned an initial weight of 200, but it is then discounted by 90 percent for a weight of 20. The stations are then scaled so the upper station has a weight of 60 percent while the lower station has a weight of 40 percent.

The river calibration has significant challenges compared to the other calibrations as well. There are fewer calibration stations than for the hydrology calibration, and those stations that do exist have far fewer observations. Generally, flow is measured daily, while water quality at a major site may have monthly samples plus storm samples. A well-monitored station may have 20–30 samples per year, while many stations have fewer samples per year or are only monitored for a few years. This data paucity relative to flow means that the same types of descriptive statistics cannot be calculated, which increases the difficulty of separating the effect of the various water quality simulation parameters. Nitrogen and phosphorus processes are linked, in that many processes affect both major constituents. This creates constraints that help to define the calibration method but also limit its flexibility.

Nitrogen, phosphorus, and sediment were all calibrated simultaneously. Relationships between simulation parameters and calibration metrics were defined and coded into software. The software automatically updated the parameters between runs similarly to the hydrology and land-based nutrient and sediment calibrations except that the sensitivities were calculated during each iteration. The sensitivities were based on the change in the calibration metric relative to the change in the parameter between the current and previous runs. The probability of interaction between variables and the likelihood of oscillation also needed to be considered. These were minimized by selecting specific calibration metrics that had minimal parameter interaction, by reducing the absolute value of the calculated sensitivity, and by constraining the calculated sensitivity to keep it within a specified range. Reasonable ranges of sensitivity were found

through sensitivity tests over all segments. It was found through trial and error that the approach of calculating, reducing, and constraining the sensitivities resulted in better calibrations than specifying a universal sensitivity as was done in the hydrology and land-based sediment and nutrient calibrations.

The calibration routine was run 30 times or until a convergence criterion was reached for each parameter. Figure 11-1 shows the calibration progress in the Rappahannock for the Phase 5.3 Model calibration. Each nitrogen loss or gain process is plotted as a separate line versus iteration. The vertical axis is percent of total nitrogen inflow that is lost through each process. Settling and denitrification are the major simulated nitrogen loss mechanisms in the Rappahannock. This particular calibration is stable after approximately 10–15 iterations. Often, there is some oscillation in the earlier iterations, which is damped out as the process continues.

Temperature was calibrated in a separate process.

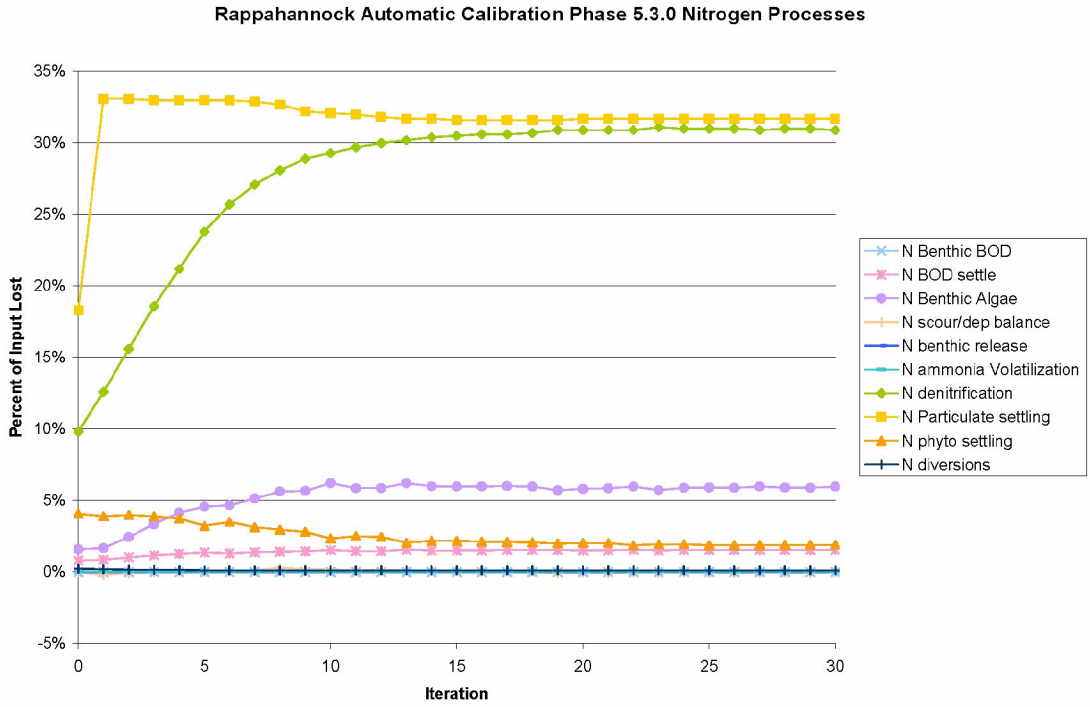

Figure 11-1. Nitrogen loss in the Rappahannock River simulation as a function of calibration iteration.

11.3.1 Calibration Metrics

The most common calibration metric is the average value of each quintile of the paired observed and simulated cumulative frequency distribution in log scale. This is illustrated in Figure 11-2. First, simulated and observed data are paired, meaning that only days in which both existed are considered. A cumulative frequency distribution is created for both. The plot is divided into five probability zones, each representing 20 percent of the values. The relevant statistic is the agreement between the average simulated value within a probability zone and the average

observed value within that zone. For example, in Figure 11-2 the observed CFD variability in total nitrogen concentration is much greater than the simulated variability as represented by the vertical frequency distribution in the simulated total nitrogen CFD. Within the calibration routine, this is represented as needing to adjust parameterization such that the upper quintile of total nitrogen is increased and the lower quintiles of total nitrogen are decreased. In the actual calibration procedure, the values are in log scale.

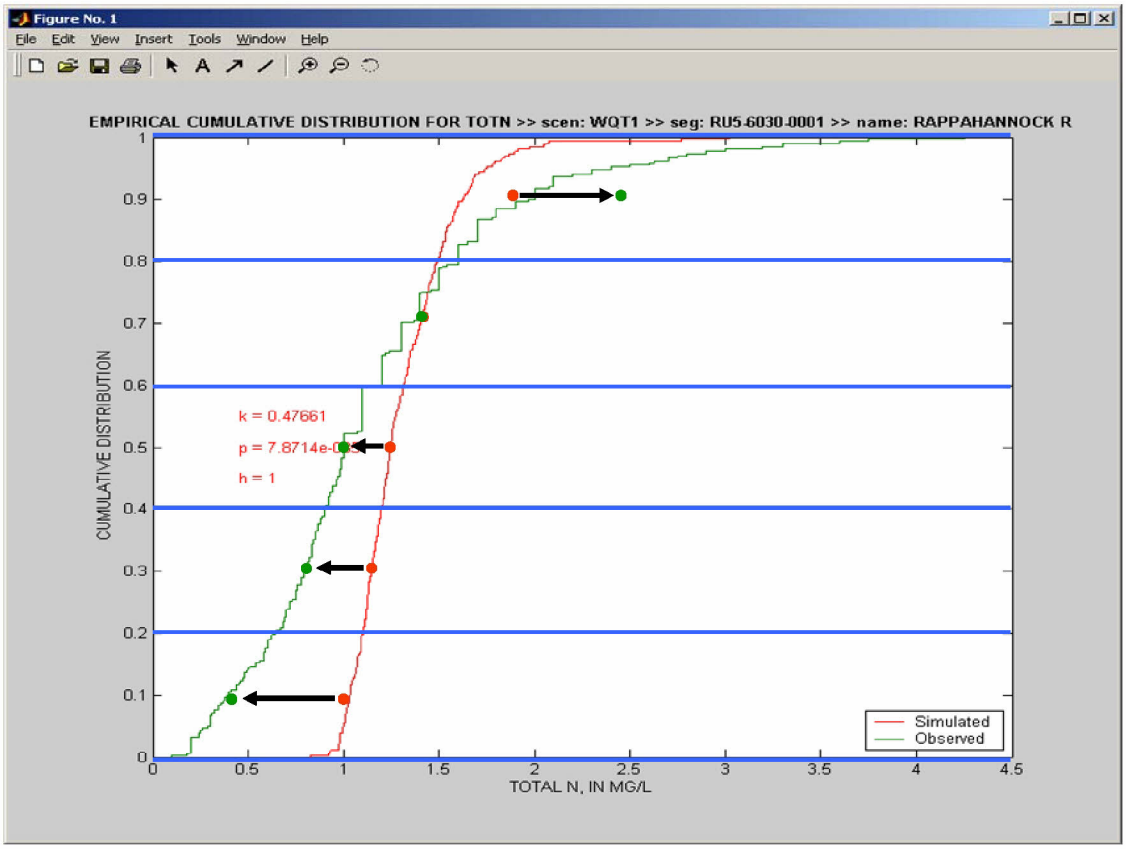

Figure 11-2. Quintiles of the paired cumulative frequency distribution.

The difference between the simulated and observed average for any quintile is referred to here as the quintile bias, where a positive quintile bias indicates that the simulated has a higher value than the observed. The lowest quintile is referred to as quintile 1 and the highest quintile is quintile 5. In Figure 11-2, quintiles 1–3 have a positive bias and quintile 5 has a negative bias. The average of all five biases is referred to as the average bias, and it is an indicator of overall bias.

Some constituents, particularly dissolved phosphate, had significant numbers of observations that were marked in the observed data set as being less than the limit of detection (LOD). For a given day with both an observed and a simulated value below the LOD, the information content of that occurrence is that there is no evidence that the simulation has to be adjusted one way or the other. For a given day with an observed value below the LOD and a simulated value greater than the LOD, the information content of that occurrence is that the simulation should be decreased at least to the LOD. For a given day with an observed value above the LOD and a simulated value

below the LOD, the information content of that occurrence is that the simulation has to increase the entire distance to the observed value. All of these occurrences are preserved by applying the following rules when constructing the cumulative frequency distributions: (1) observed values at the LOD are kept at the LOD, and (2) simulated values below the LOD are moved to the LOD value only on days when the observed value is below the LOD.

Additional information on nutrient balances and delivery factors is available in the calibration software and could be incorporated into future rules to improve calibration. Delivery factor constraints are used for the sand calibration.

In the case of temperature, model efficiency is used directly as the calibration metric.

11.3.2 Calibration Rules

Parameters that control water quality are listed in Appendix 11.A. As with hydrology, only a subset of available parameters was used in the calibration. The calibration parameters are detailed below.

Some simulated rivers have no data with which to calibrate. Most parameters for these rivers are default values, defined in Appendix 11.A. In some cases, such as for the parameters controlling chlorophyll, the parameters are calibrated to keep a water quality constituent within a specified range.

11.3.2.1 Temperature

The land parameters governing temperature were set as described in 11.2.1, and the temperature simulation in each river reach was calibrated against the next downstream gage. The simulated riverine heat balance processes include inflow, outflow, precipitation, evaporation, shortwave radiation, longwave radiation, and sensible heat transfer from the bed. The processes are mostly fully constrained, although there is a parameter associated with longwave radiation, KATRAD, which adjusts the transfer of heat from the atmosphere to the stream. The simulation of bed heat conduction is not temperature-related, so it is argued that this longwave radiation adjustment can be used to account for heat received by the stream from both longwave radiation and conduction.

Model efficiency was a concave function of KATRAD with a monotonic first derivative for all investigated river-segments. It was calibrated using a simple gradient-based optimization method.

11.3.2.2 Dissolved Oxygen

For overall mass balance, the average dissolved oxygen bias is related to the reaeration coefficient. For rivers, the reaeration coefficient is REAK; for reservoirs, it is CFOREA. To correct the shape of the distribution, the supersaturation coefficient SUPSAT is related to the fourth and fifth quintiles, while the benthic oxygen demand BENOD is related to the first and second quintiles.

11.3.2.3 Chlorophyll a

The phytoplankton settling rate, PHYSET, has the highest sensitivity to overall chlorophyll mass and affects the lower concentrations by a greater amount than the higher concentrations. PHYSET is adjusted according to the average bias for the second, third, and fourth quintiles. The

maximum algal growth rate, MALGR, is also important overall mass and especially affects higher concentrations. MALGR is related to the bias in the fifth quintile.

Observed chlorophyll values are relatively rare; only 73 stations have observations. Of those 73 stations, only a little more than half have more than 40 observations spread over the entire 21-year calibration period. The river-segments with no chlorophyll data for calibration are constrained to the interquartile range of the basinwide observed data.

11.3.2.4 Total Suspended Solids

The critical shear stresses are the most important parameters for the TSS calibration. The simulated shear stress can vary between river reaches within the same river system due to differences in the FTABLE development. Setting critical shear stress to be the same value for multiple river-segments was found to cause large differences between rivers in sediment dynamics. This issue is addressed by calculating percentiles of shear stress for each river reach after the hydrology calibration was complete. Rivers upstream of the same TSS water quality monitoring stations are calibrated together by assigning equal percentiles rather than equal shear stresses.

For example, a simulated shear stress of 0.1 pounds per square foot is reached in only 3% of the hours for the terminal segment of the Patuxent while it is reached in 25% of the hours in the next upstream segment. To deal with this issue, river-segments upstream of a particular TSS water quality monitoring station are calibrated together by assigning equal shear percentiles rather than equal shear stresses. In the example above the river-segments would be set so that they would both scour 10% of the time.

Critical shear stress for scour and critical shear stress for deposition are allowed to overlap. In a real river system of several dozen miles represented by a given reach, it is reasonable to assume that under a range of conditions, net scour and deposition would occur simultaneously at different points in the reach.

In general, silt requires higher stress to scour and can deposit under higher shear stresses than clay. Silt critical shear stresses are kept at half the distance to the 100^{th} percentile as clay. For example, if the clay critical shear stress for scour is at the 80^{th} percentile in a given river, the silt critical shear stress for scour will be set at the 90^{th} percentile.

Rivers are first checked to see if the previous iteration resulted in a complete loss of the bed sediment. In HSPF, the bed has a total mass of sediment that is available for scouring. In cases of high scour, this entire mass can be eroded, which causes a step change in the response to shear stress after the bed sediment is exhausted. If a complete loss of sediment is detected, the critical shear stresses for scour for silt and clay are raised by moving them half the distance to the 100^{th} percentile of simulated shear stress.

The critical shear stress for scour, TAUCS, changes the shape of the upper half of the cumulative frequency distribution and therefore is related in the calibration to the balance between the third and fifth quintile. If the bias is greater in the higher concentrations than near the median concentration, the critical shear stress is lowered to scour more often and vice versa. The critical

shear stress for deposition, TAUCD, is similarly related to the lower half of the distribution. Raising the critical deposition shear stress will lower the median relative to the first quintile.

The erodibility coefficient, M, has the most effect during scour events and therefore is related to the upper quintiles. The silt erodibility factor is specifically related to the fifth quintile, and the clay factor is related to the average of the fourth and fifth quintiles. Deposition velocity, W, has the most effect on lower concentrations and thus is related to the bias of the second and third quintiles. The case where the fifth quintile has a positive bias but a low erodibility coefficient indicates that the land-based loads are not being sufficiently attenuated at high flows, and therefore deposition velocity is related to the average bias rather than the low bias. Silt is kept at 10 times the clay deposition velocity.

Sand is a small part of the overall TSS observed and simulated load. Over the long term, only a small portion of the sand that enters a river reach should leave it. The coefficient on the sand load power function, KSAND, is calibrated such that only 1 percent to 10 percent of the influent sand is transported out of the river-segment.

Reservoirs are set with all critical shear stresses at the 100th percentile of simulated shear stress so that they never scour and always deposit. The calibration is controlled by the settling velocity W. The clay W is negatively related to the average bias, while the silt W is negatively related to the fourth and fifth quintile bias.

Other aspects of the sediment simulation are discussed in Section 9.

11.3.2.5 Nutrients

Nitrogen and phosphorus are connected through processes that control both nutrients. Individual species of nutrients are also connected. For example, algal uptake converts inorganic nutrients to organic nutrients for both nitrogen and phosphorus.

Settling of Refractory Organics

The settling of refractory organics is one of the more important mechanisms for attenuating excess nitrogen and phosphorus in the river reach. Ideally, the settling factor REFSET would be related to organics. If there are sufficient data, then this is the case. If there are more than twice as many total nitrogen observations as organic nitrogen observations for a station, then total nitrogen is used instead. The same applies for total phosphorus. REFSET is related in the calibration to the average for nitrogen and phosphorus of the average bias statistic.

Inorganic Nitrogen

Denitrification is related to the average bias of the lower three quintiles of nitrate concentration. The denitrification parameter, KNO320, is increased if the bias is high, but only if the total nitrogen bias is positive. Conversely, KNO320 is decreased if the bias is low, but only if the total nitrogen bias is negative.

The benthic release parameter for ammonia has a low value for use under aerobic conditions, BRTAM1, and a high value for use under anaerobic conditions, BRTAM2. The sediments are assumed anaerobic when the water column reaches the dissolved oxygen level set by the parameter ANAER. ANAER is set to roughly the 20th percentile of dissolved oxygen (DO) by

averaging the first and second quintiles of simulated DO. Ammonia is released from reservoir sediments relative to the average total nitrogen bias. These actions have the effect of releasing more ammonia from riverine sediments during warmer summer temperatures, but the amount of benthic ammonia release is guided of the calibration of the average total nitrogen bias.

Ammonium also enters the water column by attachment to sediment particles that are scoured. The concentration (mass/mass) is set by the parameters BEDNH4CLAY, BEDNH4SILT, and BEDNH4SAND. Since scoured ammonium represents a pulse of nutrients under high-flow conditions, the major effect is on the higher concentrations of nitrogen. These parameters are kept in a constant ratio to each other and are adjusted according to the fourth and fifth quintiles of total nitrogen.

The nitrification rate is used to adjust the simulated ratio of nitrate and ammonia relative to the observed ratio. The nitrification rate is increased to produce more nitrate and less ammonia, but the overall mass of inorganic nitrogen is unaffected.

Phosphate

Benthic release and scour of phosphate are handled similarly to these same processes for ammonia. A difference is that scour of phosphate is a much larger part of total phosphorus balance than the scour of ammonium is of the nitrogen balance, especially at higher concentrations. As with ammonium, scour of phosphate is related to the upper two quintiles of total phosphorus; however, benthic release in reservoirs is related to the lower four quintiles rather than the entire range. An additional mechanism to adjust the balance between dissolved phosphate and total phosphorus is the adsorption coefficient, which is related to the upper two quintiles of dissolved phosphate.

11.4 Observed Data

Observed data from 1985–2001 were received from the U.S. Geological Survey (USGS). MDE later extended the data set for the entire watershed through 2005. Additional data sets for Maryland and West Virginia were received from the Interstate Commission of the Potomac River Basin and incorporated into the data set. The following observed data sets were used for calibration with initial capitals representing the final Phase 5.3 observed data, and italics used to indicate the observed data sets which formed the final Phase 5.3 observed data.

- Total Nitrogen – Calculated as *ammonia*, plus *nitrate*, plus total *Kjeldahl nitrogen*; or calculated as *dissolved nitrogen*, plus *particulate nitrogen.*
- Nitrate – *dissolved nitrite*, plus *nitrate*; or *dissolved nitrate*
- Ammonia – *dissolved ammonia*
- Total Phosphorus – *total phosphorus*
- Dissolved Phosphate – *dissolved phosphate*
- Total Suspended Sediment – *total suspended solids*; or *total suspended sediment*
- Dissolved Oxygen – *dissolved oxygen*
- Temperature – *temperature*
- Chlorophyll *a* – *chlorophyll a*

Figure 11-3 plots the stations that were used to calibrate the Phase 5.3 Watershed Model for nitrogen and phosphorus. The calibration method that linked upstream river-segments with downstream stations and calibrated all segments in a basin simultaneously allowed for stations

that had only a few constituents. Total nitrogen, dissolved phosphate, and chlorophyll *a* were available at fewer stations than other constituents, but all stations with more than 10 observations were used in the calibration. See Table 11-1 for station totals.

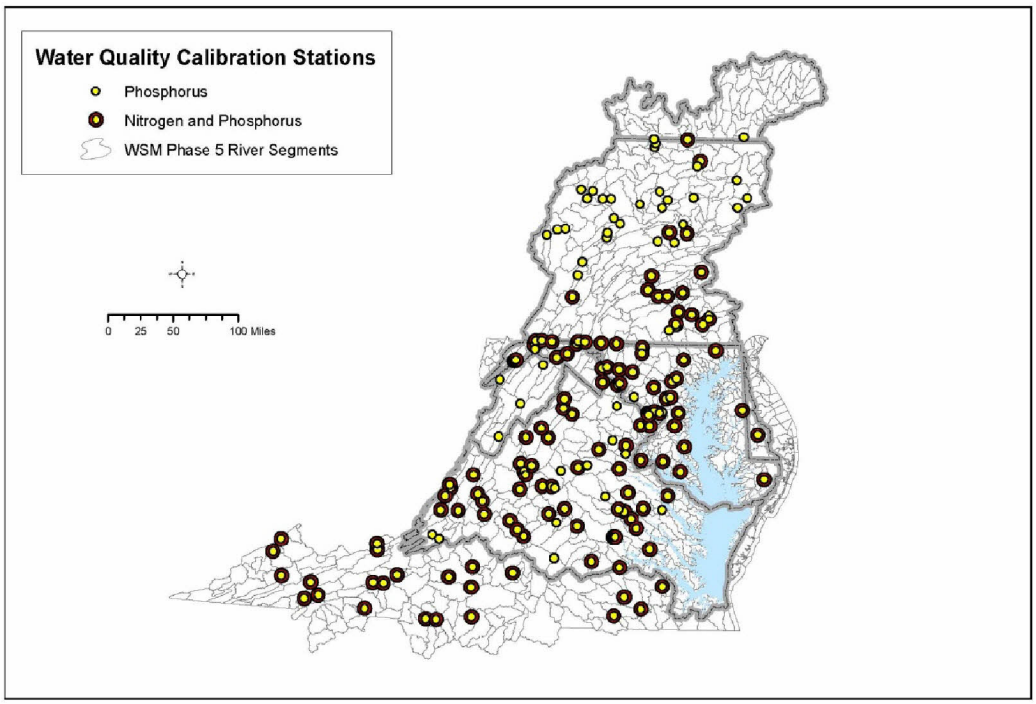

Figure 11-3. Water quality calibration stations.

Table 11-1. Number of stations used to calibrate each water quality constituent

Constituent	Stations	Constituent	Stations	Constituent	Stations
Total Nitrogen	155	Total Phosphorus	215	Dissolved Oxygen	213
Nitrate	219	Dissolved Phosphate	122	Temperature	218
Ammonia	216	Total Suspended Solids	200	Chlorophyll *a*	73

The USGS also provided annual ESTIMATOR loads for total nitrogen and total phosphorus at 30 locations in the Chesapeake Bay watershed listed below

Table 11-2. River-segment with USGS ESTIMATOR loads

EM2_3980_0001	PM2_2860_3040	PU2_3090_4050	SJ6_2130_0003	SW7_1640_0003
JA5_7480_0001	PM4_4040_0003	PU3_3290_3390	SL3_2420_2700	XU2_4330_4480
JL6_7430_7320	PM7_4820_0001	PU3_3680_3890	SL9_2490_2520	XU3_4650_0001
JL7_6800_7070	PS2_6730_6660	PU6_4020_3870	SL9_2720_0001	YM4_6620_0003

JL7_7100_7030	PS3_5100_5080	RU2_5940_6200	SU7_0850_0730	YP3_6330_6700
PL0_4510_0001	PS5_5240_5200	RU5_6030_0001	SU8_1610_1530	YP4_6720_6750

11.5 Regional Factors

As described in Section 10, land use simulations are calibrated to target nitrogen and phosphorus loads based on an overall nutrient balance calculation. The target loads are centered on literature-based medians. In a natural system, physical processes might create an environment that is particularly conducive to transport of one or both types of nutrients. An example would be a land-segment with soil having a high groundwater denitrification potential. These physical characteristics are not simulated explicitly in the Watershed Model.

Without simulation of different potential delivery in the land use simulation, the gains or losses could be made up in the river system in order to meet water quality measurements in the rivers. Figure 11-4 is a chart of the relative importance of nitrogen and phosphorus losses in the calibrated Patuxent River above Bowie, Maryland, which has several significant reservoirs. The processes represented by checkered symbols are those that are controlled by separate parameters for nitrogen and phosphorus. The others, which represent approximately 90 percent of the attenuation, are controlled by the same parameters for both nutrients. This makes it difficult to correct a large imbalance between nitrogen and phosphorus within the river calibration, but even if it were possible, it would likely not be the correct method for modeling the losses if they land-based processes were causing the imbalance.

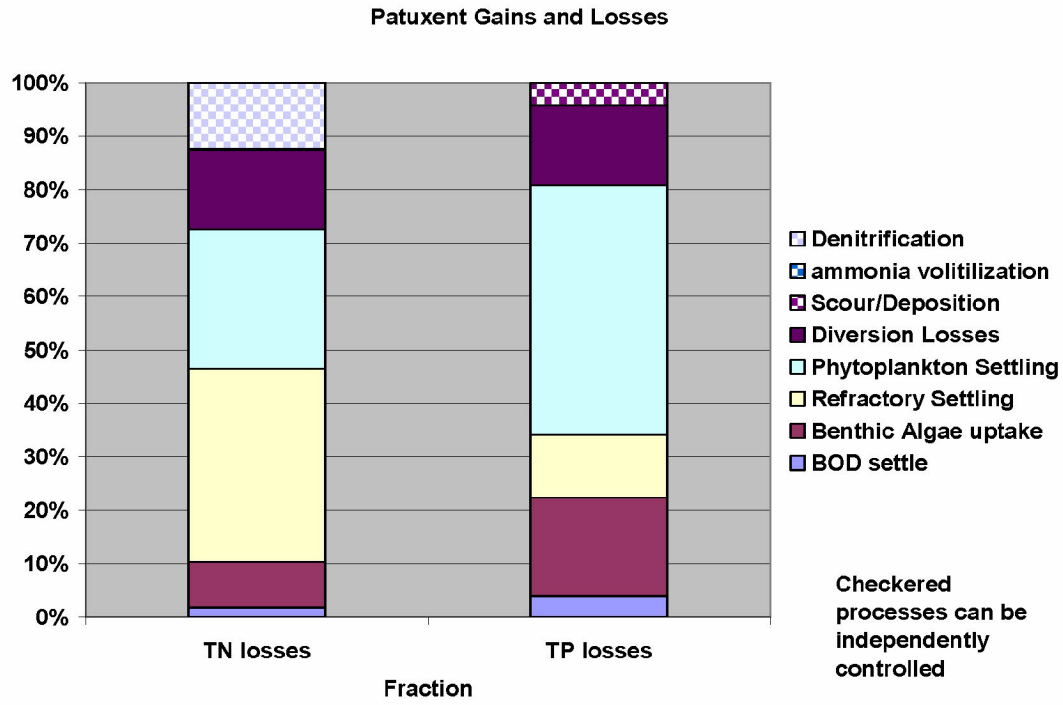

Figure 11-4. Gains and losses in the Patuxent River simulation.

Without *a priori* information on the relative land-based loads from different geologies and without a method of correcting large imbalances in the river calibration, it was necessary to adjust land-based loads to reach a satisfactory calibration of nutrients in the rivers. In a typical smaller-scale watershed calibration to a single monitoring station, a modeler can typically make adjustments to nutrient loads consistent with available local information and literature values. Given the large number of calibration stations and the complexity of the model, it was infeasible to adjust the land simulation of nutrients to obtain a better river calibration, so the EoS land use loads were modified by *regional factors* to achieve a better nutrient calibration in rivers and an automated procedure to calculate the regional factors was implemented. Edge-of-stream loads from the land simulation were multiplied by the regional factors before they are input into river reaches. These regional factors were applied to all land-based loads upstream of a given site.

Initial efforts using *a priori* estimates of regional factors using the USGS ESTIMATOR (USGS 2007), and SPARROW models failed to produce satisfactory results likely due to combined errors from those two models, Phase 5, and unaccounted-for regional biases, which are not necessarily addressed in SPARROW. Correcting using the median concentration bias for each station worked reasonably well, but it caused large swings in regional factors in areas that were reasonably close geographically and geologically. The final method for Phase 5.3 adjusting nitrogen and phosphorus loads from each land use by a factor not to exceed 2.0 and not to be less than 0.5 based on the agreement between a calibration as described above and the USGS ESTIMATOR loads at the 30 segments listed in Table 11-2. A mass balance is performed around a sub-watershed, which is defined as the watershed between an ESTIMATOR site and any upstream ESTIMATOR site. The regional factor is applied to land-based sources in order to best match the long-term load at the site.

Regional factors are also estimated for the ungaged areas near the tidal water, where there is only one ESTIMATOR site and much of the area is unmonitored. For the Coastal Plain, regional factors are estimated based on the concentration difference between simulated and observed at 14 sites for nitrogen and 18 sites for phosphorus. The estimated regional factors are then applied to the rest of the Coastal Plain based on observed patterns in the regional factors. Regional factors are grouped by region into Eastern Shore, Upper Western Shore, and Lower Western Shore, and also by hydrogeomorphic region. A number of Kolmogorov-Smirnov tests are used to determine the likelihood that a true difference in the distribution between regions exists. For Phase 5.3 the strongest difference was between various Coastal Plain geologies and the combination of Piedmont and Mesozoic lowlands geologies. The Coastal Plain was given average regional factors of 0.6 for nitrogen and 1.2 for phosphorus, while the Piedmont and Mesozoic lowlands were given factors of 1.0 for nitrogen and 0.65 for phosphorus.

11.6 Calibration Results

The calibration metrics and validation approach are based on recommendations from the 2008 Scientific and Technical Advisory Committee (STAC) review of the Watershed Model (STAC 2008). Calibration metrics presented in this document and recommended by the STAC review are meant to give an overview of model performance across stations rather than a detailed look at a particular station. Plots of simulated and observed instantaneous concentrations and loads are available on the calibration website,

ftp://ftp.chesapeakebay.net/Modeling/phase5/Phase%205.3%20Calibration/Calibration_pdf/all_validation.pdf

Figures 11-5 through 11-7 display the agreement between three models of load at the river-segments with ESTIMATOR model data available. Note that the data are log scale. ESTIMATOR is an empirical model of load based on observed data at a point in a river (USGS 2007). Two phases of the Watershed Model are shown on these figures. The Phase 5.3 Community Watershed Model is the model used in development of the Chesapeake 2010 TMDL. Phase 5.2 was a development version that is included to show robustness of the calibration procedures. The black line is the 1:1 plot for ESTIMATOR. In general, the agreement among all three models is very good. The result is unsurprising since ESTIMATOR is explicitly used in the calibration of both Phase 5 versions. The increased variability of phosphorus over nitrogen and of sediment over phosphorus is indicative of the increased uncertainty for those constituents for all three models.

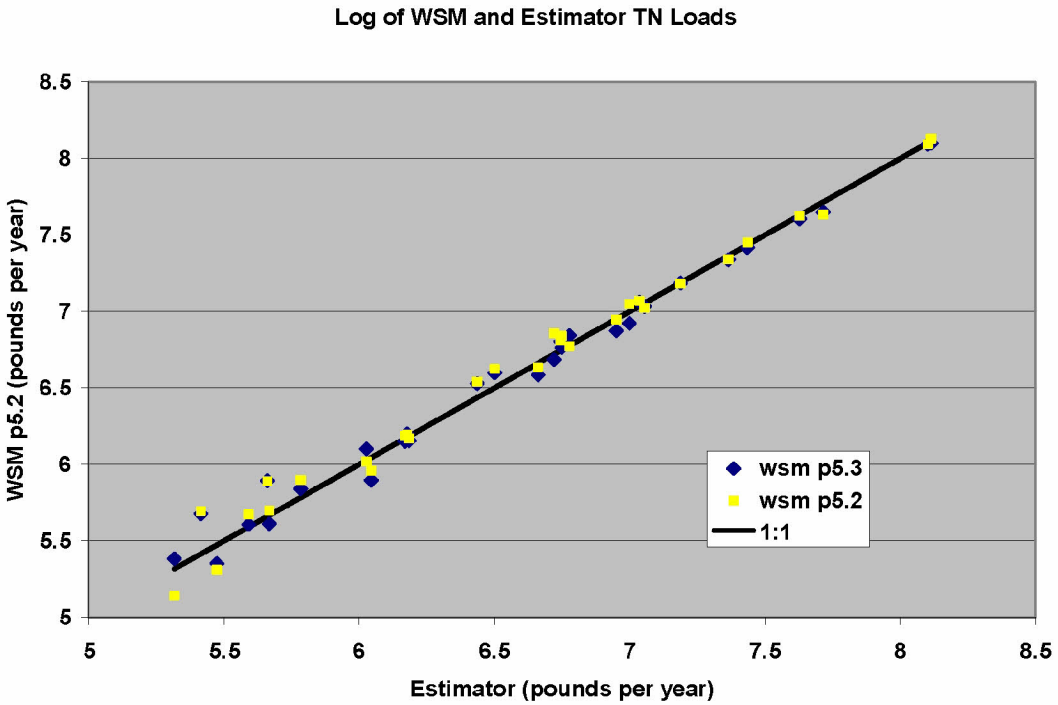

Figure 11-5. ESTIMATOR and Watershed Model nitrogen loads.

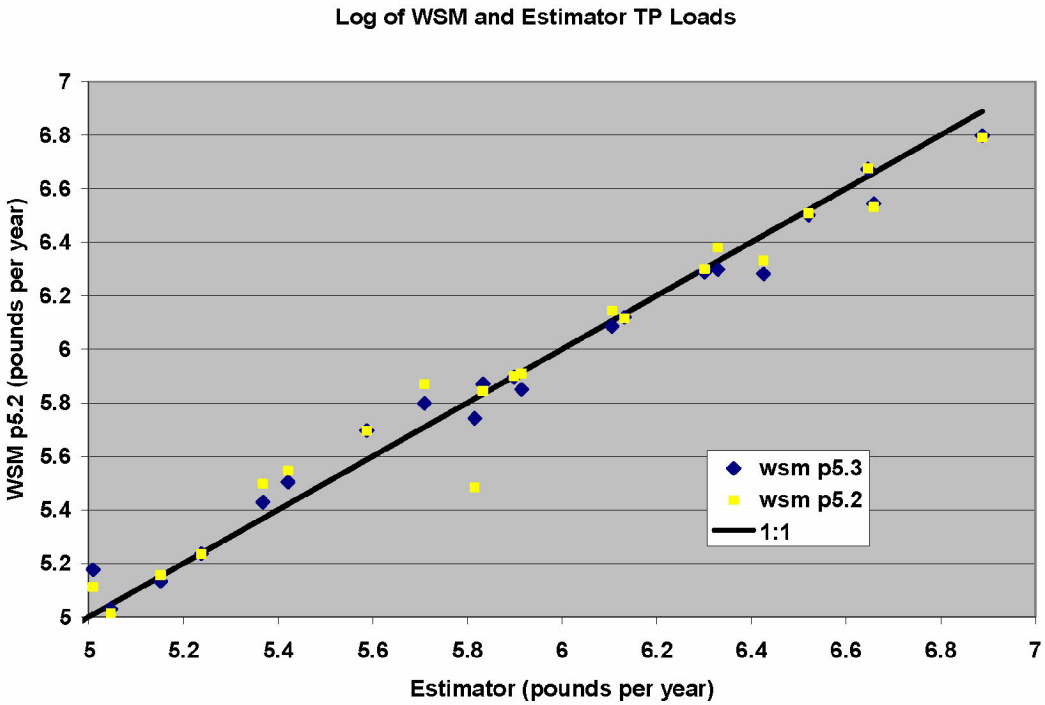

Figure 11-6. ESTIMATOR and Watershed Model phosphorus loads.

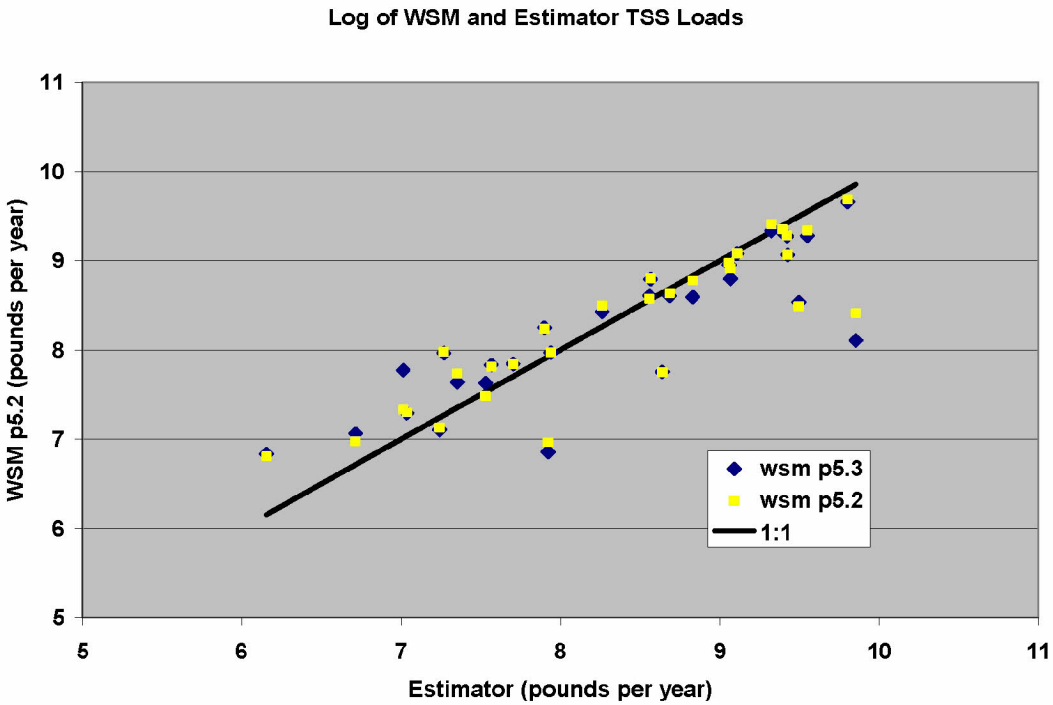

Figure 11-7. ESTIMATOR and Watershed Model sediment loads.

Figures 11-8 through 11-10 show the correlation of the annual ESTIMATOR loads and the annual Watershed Model loads at the nine river input stations. Correlation is an indicator of temporal consistency between different models of load. It is again not surprising that the correlations are high given that the flow, which is a powerful determinant of load, is well calibrated in the Phase 5.3 model and observed flow is used in ESTIMATOR regressions. Even though phosphorus and sediment are generally more difficult to calibrate than nitrogen, the correlation shows up as comparable to nitrogen because there is more variability to describe.

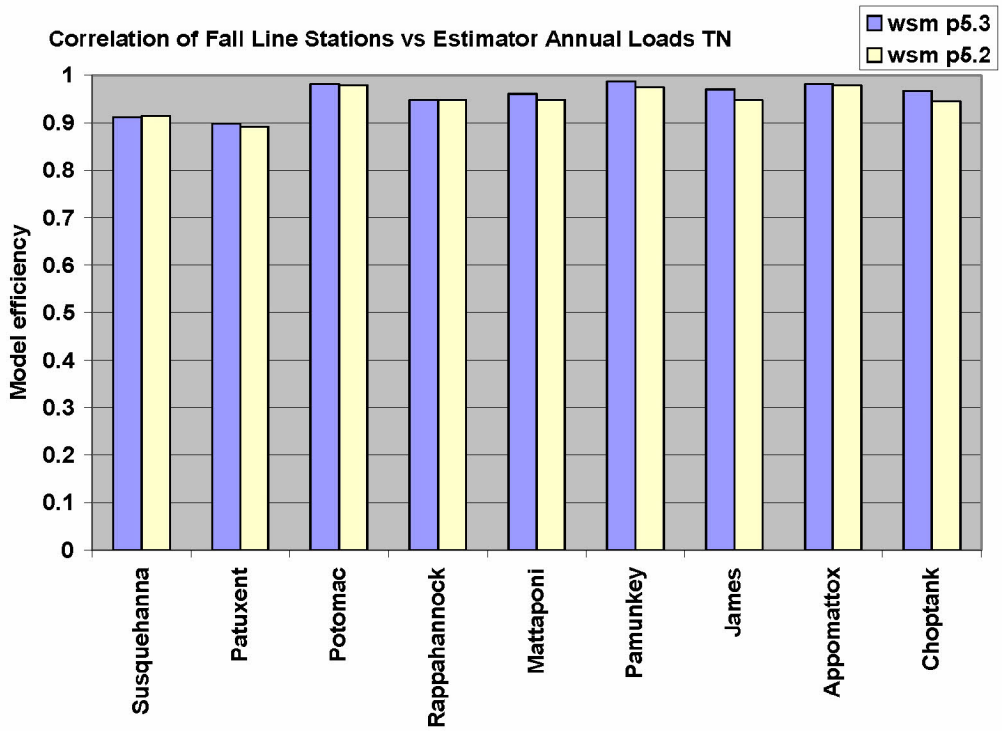

Figure 11-8. ESTIMATOR and Watershed Model annual nitrogen load correlation.

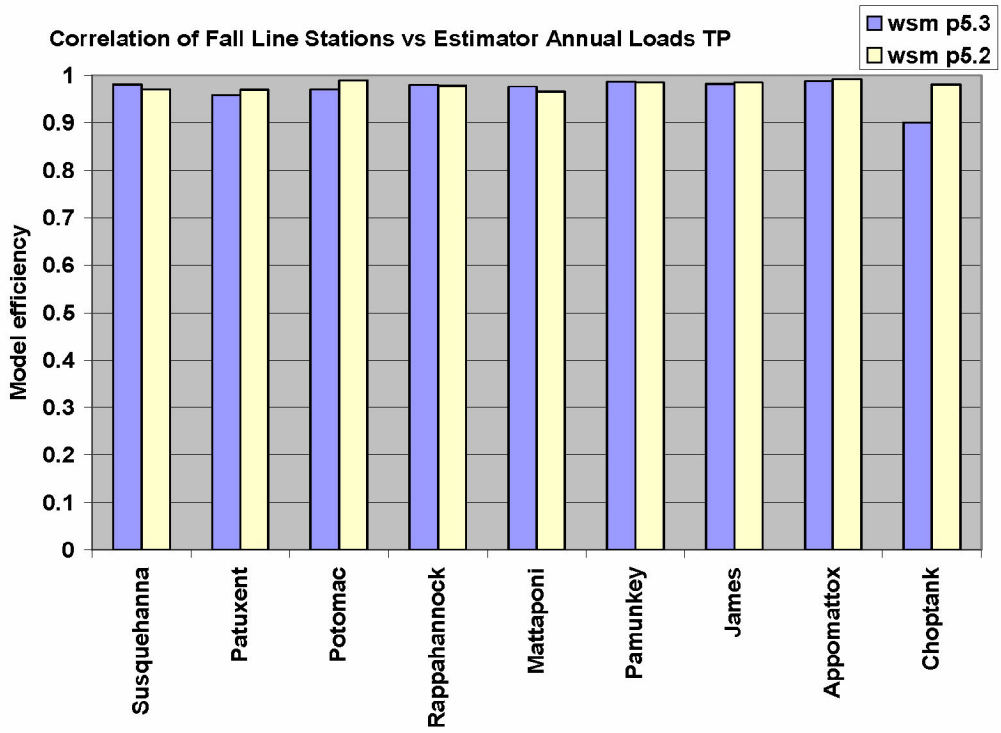

Figure 11-9. ESTIMATOR and Watershed Model annual phosphorus load correlation.

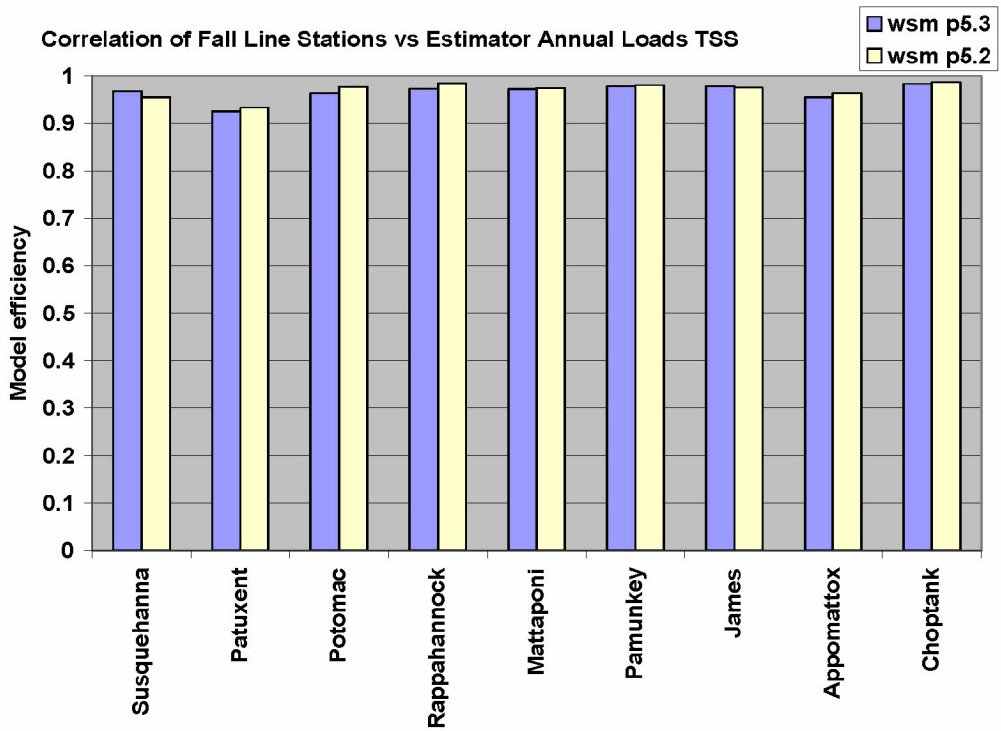

Figure 11-10. ESTIMATOR and Watershed Model annual sediment load correlation.

Figures 11-11 through 11-13 are representations of the distribution of concentration. For the observed, the unbiased concentrations from USGS (2007) were used. The stations selected are the stations in the referenced report. The numbers in the key represent percentiles of the distribution. The five lines represent, from top to bottom, the 90th, 75th, 50th, 25th, and 10th percentiles of the distribution. The stations are ordered by the USGS median value. The plot is meant as an indicator of the watershed model matching the observed spatial variability. All three figures show that the watershed model matches the spatial variability of the observed concentration data both in the median value and in the distribution. Figure 11-13 for sediment shows a difference in the lower percentiles; however, it is clear from the overlapping observed lines that this is due to censored sediment data.

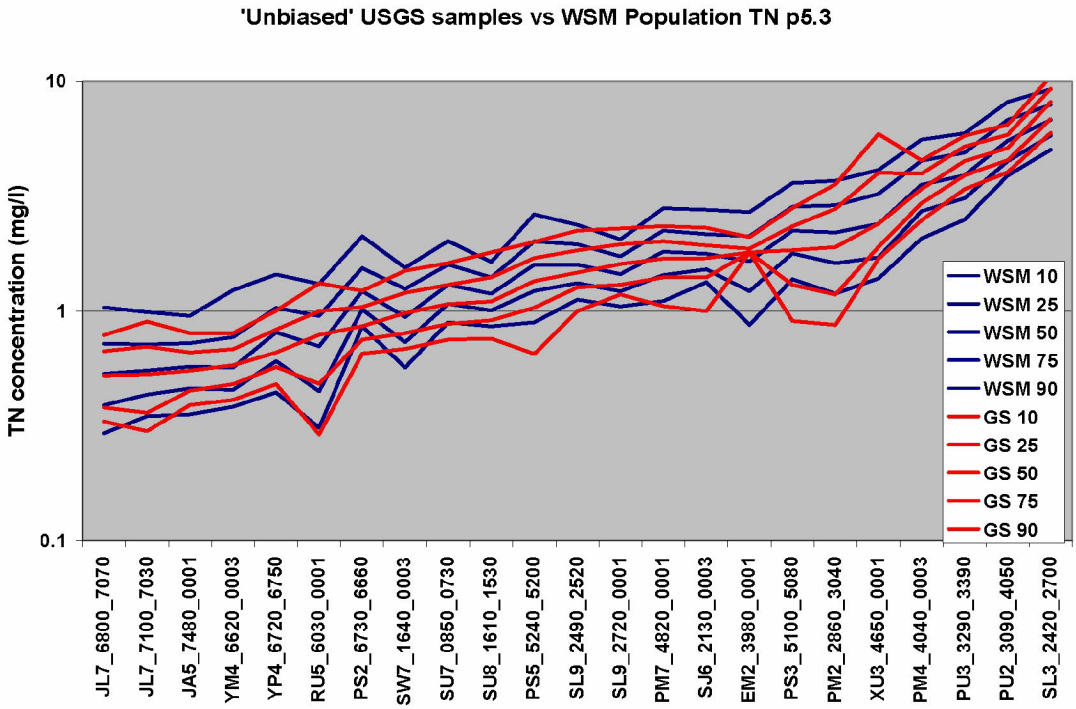

Figure 11-11. Unbiased USGS and Watershed Model nitrogen concentrations.

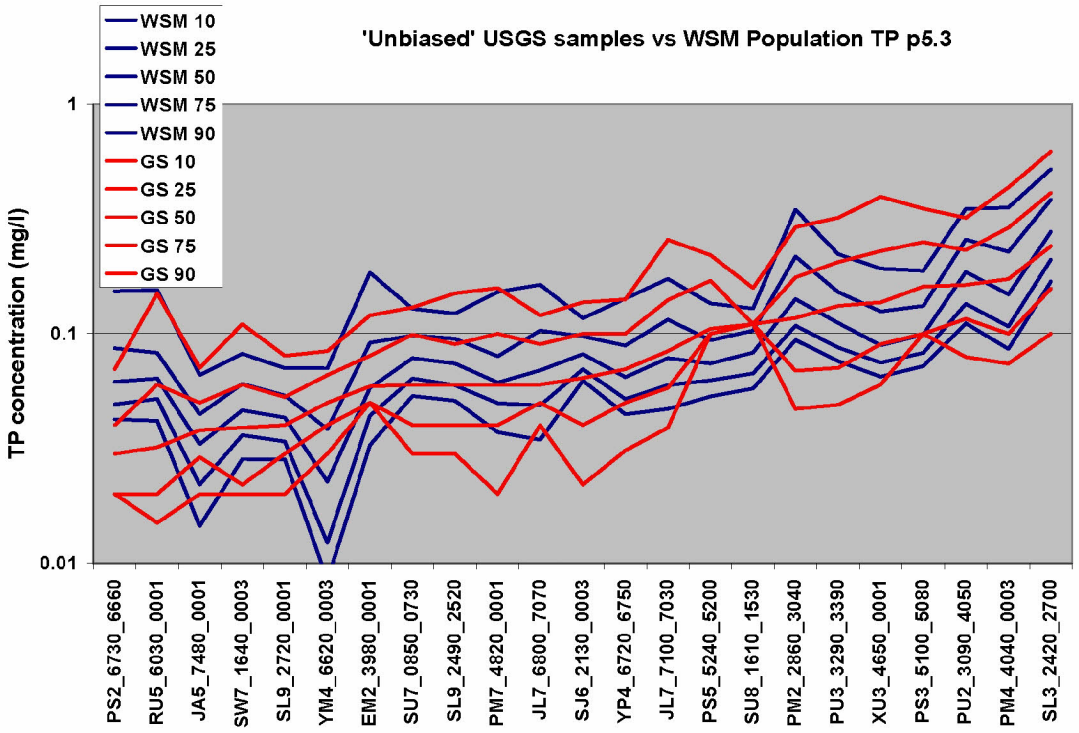

Figure 11-12. Unbiased USGS and Watershed Model phosphorus concentrations.

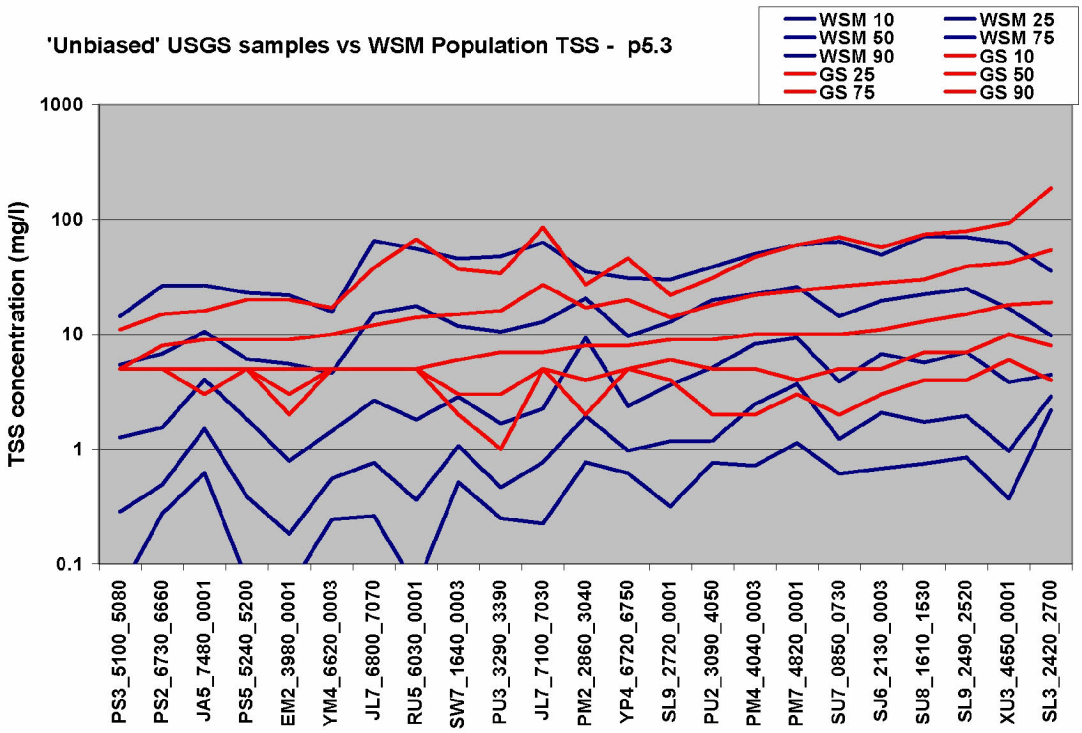

Figure 11-13. Unbiased USGS and Watershed Model sediment concentrations.

A subset of the 33 calibration sites with adequate data for validation was selected. The river-segments containing these stations are shown in Table 11-3. Validation data sets were compiled from the most recent 25 percent of the observations for each selected station, with the extra provision that observations between 1991 and 2000 are held within the calibration data set. This extra provision is to allow for the most accurate calibration for the selected 10-year scenario hydrologic period that will be used for the Bay TMDL. If insufficient observations can be found after 2000, observations from the 1985–1990 period are used for validation. If there are still too few observations for validation, observations within the 1991–2000 period are withheld.

Table 11-3. River-segments containing validation monitoring stations

EM2_3980_0001	PM2_2860_3040	PU6_3752_4080	SL9_2490_2520	WU2_3020_3320
JA5_7480_0001	PM7_4820_0001	RU5_6030_0001	SL9_2720_0001	XU0_4130_4070
JL7_7100_7030	PU2_3090_4050	SL1_2390_2420	SU7_0850_0730	XL1_4690_0001
PL0_4510_0001	PU3_3680_3890	SL3_2290_2260	SU8_1610_1530	XU3_4650_0001
PL1_4460_4780	PU3_3290_3390	SL3_2420_2700	WM1_3660_3910	XU2_4330_4480
PM1_3510_4000	PU6_3610_3530	SL3_2460_2430	WM3_3880_4060	YP4_6720_6750
PM1_3120_3400	PU6_4020_3870	SL4_2140_2240		

Figures 11-14 through 11-16 are the validation plots. Each figure plots a statistic indicative of the goodness of fit between observed and simulated concentrations for the calibration time period and the validation time period at the representative stations in Table 11-3. The goal of the validation is to demonstrate that the parameters determined in the calibration can successfully simulate observed data that were not used to calibrate the model. The calibration is successfully validated if the calibration statistics are not consistently better than the validation statistics.

Figure 11-14 is a comparison of the calibration and validation for nitrogen. The relevant statistic is the Kolmogorov-Smirnov (K-S) test statistic. This is the maximum probability difference between the simulated cumulative frequency distribution (CFD) and the observed CFD. It is an indicator of the probability that the two samples are selected from the same distribution. A lower value means a better match. Any point above the 1:1 line means that the calibration had a better test statistic than the validation. As shown in Figure 11-14, the distributions above and below the 1:1 line are similar, indicating a fully successful validation.

For phosphorus and sediment, the K-S statistic was not necessarily as indicative of a good match between the distribution of observed and simulated concentrations. Phosphorus and sediment loads are governed primarily by the highest concentrations, which generally occurs at the highest flows. The bias of the mean value of the upper quintile of the simulated and observed CFD is therefore used as the relevant statistic. Figures 11-15 and 11-16 show the calibration and validation statistics plotted against each other. The extent to which the points plot near the 1:1 line or 1:-1 line or are evenly distributed throughout the space is an indicator of a full validation. Points along the vertical axis are an indicator of a failed validation. In both cases, more points are indicative of better agreement in the calibration than the validation, indicating that there is some calibration effect. However, there is a reasonable distribution of points throughout the space, indicating some level of success in the validation.

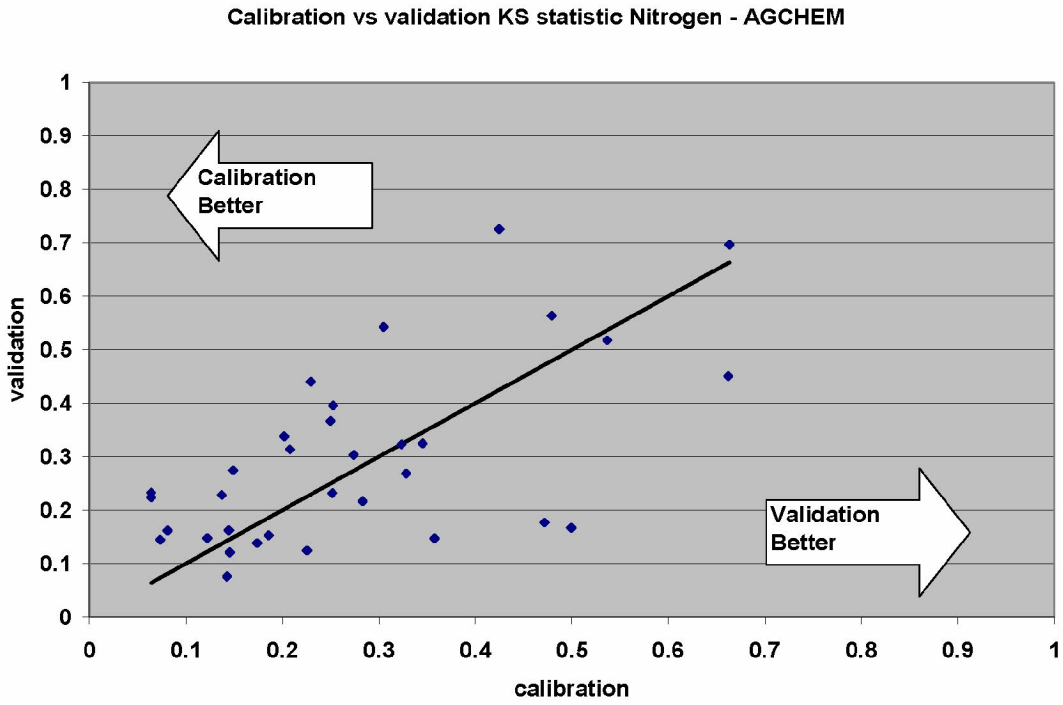

Figure 11-14. Nitrogen validation.

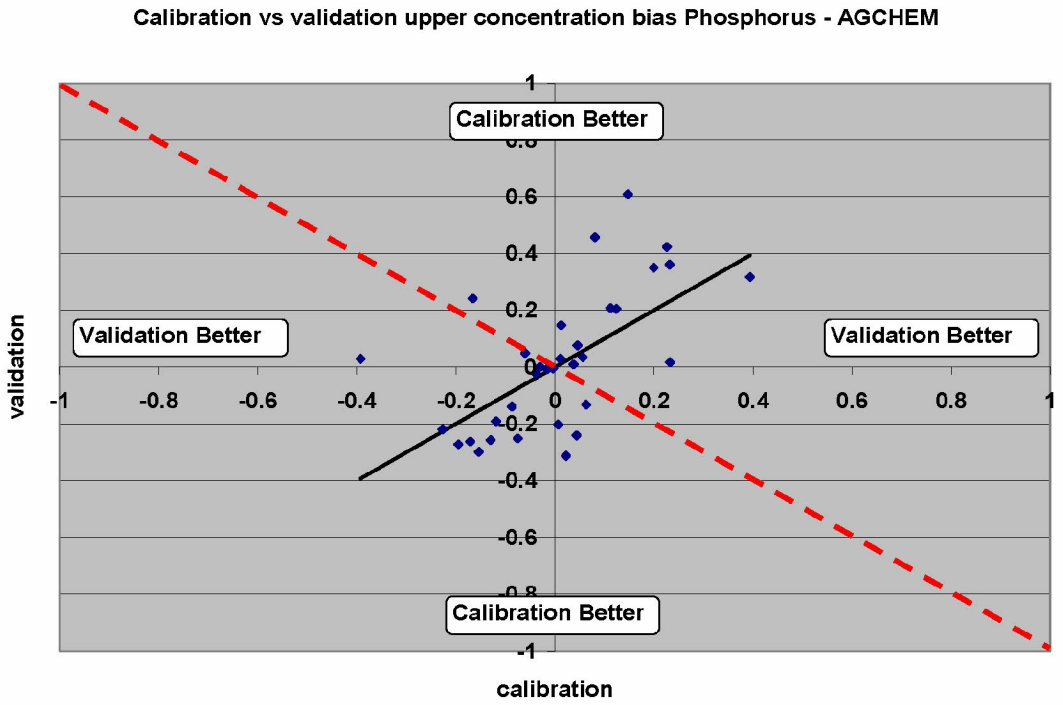

Figure 11-15. Phosphorus validation.

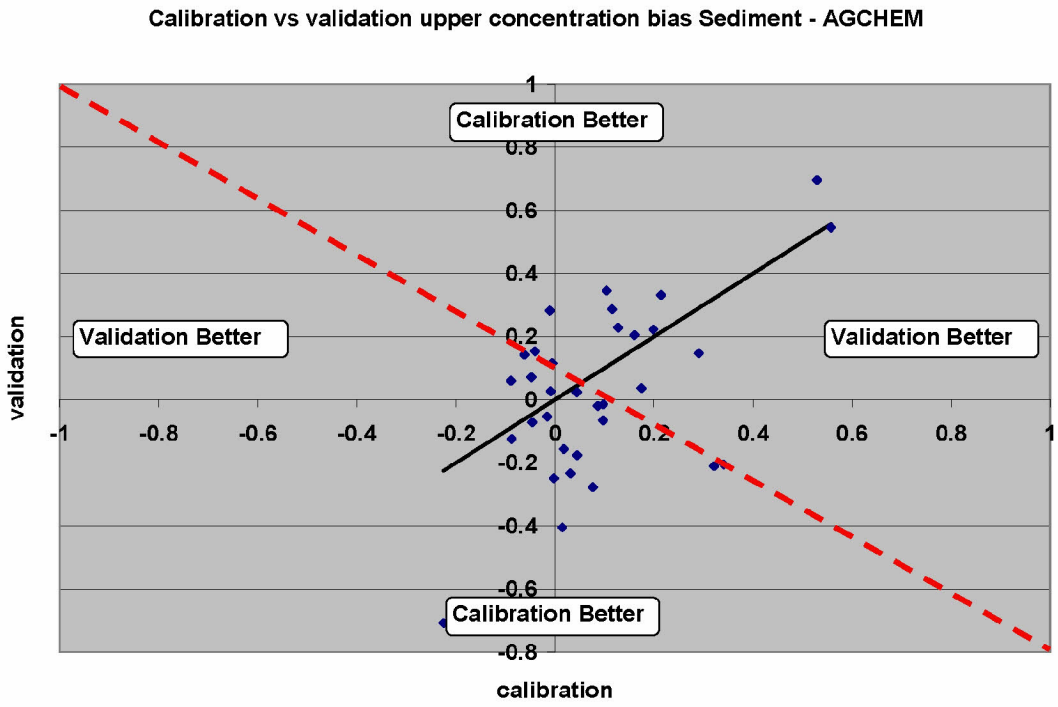

Figure 11-16. Sediment validation.

Edge-of-Stream and Delivered Loads

The Phase 5.3 EoS and delivered loads can be downloaded by CVS files (comma separated value files) which contain EoS and delivered loads from the Phase 5.3 Watershed Model calibration. The loads of different constituents are displayed by land river segment and land-use. Edge-of-stream loads are abbreviated as eos, and delivered loads as del.

The data is organized as follows:
Land Segment, River Segment, Land Use, Edge of Stream/Delivered, Constituent, Load

All loads are supplied as averages over the calibration period in pounds (lbs), except for the heat term, which is in BTUs. Acres are provided so that pounds per acre can be calculated for comparison across land uses or land/river segments if desired.

To access loads specific to a model land-segment or river-segment, use these Excel files to determine the model segments of interest.
http://ches.communitymodeling.org/models/CBPhase5/datalibrary/landnames.xls
http://ches.communitymodeling.org/models/CBPhase5/datalibrary/rivernames.xls

To find load data specific to a land use category for any segment use Table 11.4 as guidance and the URL below. Note that the file will be truncated if opened in MS Excel because the file is 75 MB in size.

ftp://ftp.chesapeakebay.net/Modeling/phase5/Phase 5.3 Calibration/Model Output/AllLoads_watershed_1991_2000_p53cal.csv

For additional information on the model land uses, consult the Section 4 - Land Use document at:
http://ches.communitymodeling.org/models/CBPhase5/documentation.php#p5modeldoc

Table 11-4. Guidance for land use abbreviations in the detailed land-segment load file

Model Land Uses	Phase 5 Watershed Model	Abbreviation
Forest		
	Forest, Woodlots & Wooded	for
	Harvested Forest	hvf
Agriculture		
	High Tillage with Manure	hwm
	High Tillage without Manure	hom
	Low Tillage with Manure	lwm
	Nutrient Management High Tillage with Manure	nhi
	Nutrient Management High Tillage w/o manure	nho
	Nutrient Management Low Tillage	nlo
	Hay Land	
	Alfalfa	alf
	Nutrient Management alfalfa	nal
	Hay With Nutrients	hyw
	Hay w/o Nutrients	hyo

Detailed calibration information at all of the individual monitoring stations used can be found by river basin using the links below.

Susquehanna
ftp://ftp.chesapeakebay.net/Modeling/phase5/calibration_pdfs/p53_2010_02/SC.pdf

Potomac
ftp://ftp.chesapeakebay.net/Modeling/phase5/calibration_pdfs/p53_2010_02/PM.pdf

Patuxent
ftp://ftp.chesapeakebay.net/Modeling/phase5/calibration_pdfs/p53_2010_02/XU.pdf

Rappahannock
ftp://ftp.chesapeakebay.net/Modeling/phase5/calibration_pdfs/p53_2010_02/RU.pdf

York - Pamunkey
ftp://ftp.chesapeakebay.net/Modeling/phase5/calibration_pdfs/p53_2010_02/YP.pdf

York - Mattaponi
ftp://ftp.chesapeakebay.net/Modeling/phase5/calibration_pdfs/p53_2010_02/YM.pdf

James
ftp://ftp.chesapeakebay.net/Modeling/phase5/calibration_pdfs/p53_2010_02/JL.pdf

Appomattox
ftp://ftp.chesapeakebay.net/Modeling/phase5/calibration_pdfs/p53_2010_02/JA.pdf

Coastal Plain
ftp://ftp.chesapeakebay.net/Modeling/phase5/calibration_pdfs/p53_2010_02/CP.pdf

All Fall Line Stations
ftp://ftp.chesapeakebay.net/Modeling/phase5/calibration_pdfs/p53_2010_02/p5_fall_lines.pdf

Calibration Stations Outside Chesapeake Watershed
ftp://ftp.chesapeakebay.net/Modeling/phase5/calibration_pdfs/p53_2010_02/GY.pdf

References

Baier, G.T., T. Cohn, and E. Gilroy. 1995. Instructions for Using the Estimator Software. <http://www159.pair.com/cohns/TimCohn/TAC_Software/Estimator/Est_user_man04.html> Accessed December 30, 2010.

Bicknell, B.R., J.C. Imhoff, J.L. Kittle, Jr., A.S. Donigian, Jr., and R. C. Johnson. 1997. Hydrological Simulation Program - Fortran (HSPF): User's Manual for Release 11. U.S. Environmental Protection Agency. EPA/600/R-97/080, 755 p. Athens, GA.

Bicknell, B.R., J. C. Imhoff, J. L. Kittle, Jr., T. H. Jobes, and A. S. Donigian, Jr. 2001. *Hydrologic Simulation Program – Fortran, HSPF Version 12 Users Manual.* U. S. Environmental Protection Agency, Office of Research and Development, Athens, GA. USGS Hydrologic Analysis Software Support Program, Reston, VA.

Cohn, T., L.L. DeLong, E.J. Gilroy, R.M. Hirsch, and D.K. Wells. 1989. Estimating Constituent Loads. *Water Resources Research*, 25(5): 937-942.

Cohn, T., D.L. Caulder, E.J. Gilroy, L.D. Zynjuk, and R.M. Summers. 1992. The Validity of a Simple Statistical Model for Estimating Fluvial Constituent Loads: An Empirical Study Involving Nutrient Loads Entering Chesapeake Bay. *Water Resources Research*, 28(9): 2353-2364.

Cohn, T. 2002. Estimator 2002, A Beta Release. <http://www159.pair.com/cohns/TimCohn/TAC_Software/Estimator/e2002/index.html>. Accessed December 30, 2010.

Donigian, A.S., Jr., Imhoff, J.C., Bicknell, Brian, Kittle, J.L., Jr., 1984. Application guide for Hydrological Simulation Program--Fortran (HSPF): U.S. Environmental Protection Agency, Environmental Research Laboratory, EPA-600/3-84-065, 177 p. Athens, GA.

Johanson, R.C., Imhoff, J.D., and Davis, H.H., Jr., 1980. Users manual for hydrological simulation program - Fortran (HSPF): Environmental Research Laboratory, EPA-600/9-80-015, April 1980. Athens, GA.

(STAC) Scientific and Technical Advisory Committee. 2008. *Second Review of the Phase 5 Community Chesapeake Bay Watershed Model* STAC Publication Number: 08-003. Edgewater, MD. February 20, 2008. <http://archive.chesapeakebay.net/pubs/subcommittee/mdsc/Review_Chesapeake_Bay_Watershed_Modeling_Effort_2008.pdf>. Accessed December 30, 2010.

USGS (U.S. Geological Survey). 2006. *Changes in Streamflow and Water Quality in Selected Nontidal Basins in the Chesapeake Bay Watershed, 1985–2004*. Scientific Investigations Report 2006-5178. U.S. Geological Survey, Reston, VA.

Cite as:

USEPA (U.S. Environmental Protection Agency). 2010. *Chesapeake Bay Phase 5.3 Community Watershed Model*. EPA 903S10002 - CBP/TRS-303-10. U.S. Environmental Protection Agency, Chesapeake Bay Program Office, Annapolis MD. December 2010.

Errata Revision March 9, 2011

Appendix A: Water Quality Parameters

Lake: Parameter applies to:
- 2 = both rivers and reservoirs
- 1 = reservoirs only
- 0 = rivers only

Module, Table, Parameter: HSPF module, variable table, and parameter name. See HSPF documentation for description.

Start: Starting value in the calibration. For uncalibrated parameters, this is the final value as well.

Min and Max: Enforced minimum and maximum for the calibration

Type: Refers to calibration update sensitivity.
- A = Additive. The parameter is somewhat linearly related to the calibration metric. Sensitivities are calculated and adjustments made on the untransformed parameter.
- M = Multiplicative. The log of the parameter is somewhat linearly related to the calibration metric. Sensitivities are calculated and adjustments made on the log transformed parameter.

Lake	Module	Table	Parameter	Start	Min	Max	Type
2	ADCALC	ADCALC-DATA	CRRAT	1.5	1	2	A
2	SEDTRN	SAND-PM	KSAND	0.001	0.00001	1	M
2	SEDTRN	SAND-PM	EXPSND	4	2	5	A
2	SEDTRN	SILT-CLAY-PM#1	SW	0.01	0.0001	0.1	M
2	SEDTRN	SILT-CLAY-PM#1	STAUCD	95	10	100	A
2	SEDTRN	SILT-CLAY-PM#1	STAUCS	95	50	100	A
2	SEDTRN	SILT-CLAY-PM#1	SM	0.1	0.001	10	M
2	SEDTRN	SILT-CLAY-PM#2	CW	0.001	0.00001	0.01	M
2	SEDTRN	SILT-CLAY-PM#2	CTAUCD	90	10	100	A
2	SEDTRN	SILT-CLAY-PM#2	CTAUCS	90	50	100	A
2	SEDTRN	SILT-CLAY-PM#2	CM	0.1	0.001	10	M
2	RQUAL	SCOUR-PARMS	SCRVEL	5	1	10	A
2	RQUAL	SCOUR-PARMS	SCRMUL	2	1	5	A
2	OXRX	OX-GENPARM	KBOD20	0.05	0.005	0.5	M
2	OXRX	OX-GENPARM	KODSET	0.01	0.001	0.1	M
2	OXRX	OX-GENPARM	SUPSAT	1.15	1	1.35	A
2	OXRX	OX-BENPARM	BENOD	40	1	200	A

2	OXRX	OX-BENPARM	BRBOD1	0.01	0.001	0.1	M
2	OXRX	OX-BENPARM	BRBOD2	0.01	0.001	0.1	M
0	OXRX	OX-REAPARM	REAK	0.5	0.05	5	M
1	OXRX	OX-CFOREA	CFOREA	1	0.1	2	A
1	NUTRX	NUT-BENPARM	BRTAM1	0.5	0.01	40	A
1	NUTRX	NUT-BENPARM	BRTAM2	2	0.04	100	A
1	NUTRX	NUT-BENPARM	BRPO41	0.05	0.001	2	A
1	NUTRX	NUT-BENPARM	BRPO42	0.25	0.005	10	A
1	NUTRX	NUT-BENPARM	ANAER	1	0.2	7	A
2	NUTRX	NUT-NITDENIT	KTAM20	0.2	0.005	1	M
2	NUTRX	NUT-NITDENIT	KNO320	0.01	0.001	0.1	M
2	NUTRX	NUT-NITDENIT	DENOXT	20	10	30	A
2	NUTRX	NUT-NH3VOLAT	EXPNVG	0.5	0.3	0.7	A
2	NUTRX	NUT-NH3VOLAT	EXPNVL	0.6667	0.4	0.9	A
2	NUTRX	NUT-BEDCONC	NH4-sand	2	0.2	20	M
2	NUTRX	NUT-BEDCONC	NH4-silt	20	2	200	M
2	NUTRX	NUT-BEDCONC	NH4-clay	200	20	2000	M
2	NUTRX	NUT-BEDCONC	PO4-sand	2	0.2	10	M
2	NUTRX	NUT-BEDCONC	PO4-silt	20	2	100	M
2	NUTRX	NUT-BEDCONC	PO4-clay	200	20	1000	M
2	NUTRX	NUT-ADSPARM	NH4-sand	150	1.5	1500	M
2	NUTRX	NUT-ADSPARM	NH4-silt	1500	15	10000	M
2	NUTRX	NUT-ADSPARM	NH4-clay	15000	150	15000	M
2	NUTRX	NUT-ADSPARM	PO4-sand	3333	33	33333	M
2	NUTRX	NUT-ADSPARM	PO4-silt	10000	100	100000	M
2	NUTRX	NUT-ADSPARM	PO4-clay	30000	300	300000	M
2	PLANK	PLNK-PARM1	NONREF	0.75	0.5	0.9	A
2	PLANK	PLNK-PARM1	LITSED	0.01	0.001	1	M
2	PLANK	PLNK-PARM1	ALNPR	0.8	0.6	0.9	A
2	PLANK	PLNK-PARM1	EXTB	0.1	0.01	1	M
2	PLANK	PLNK-PARM1	MALGR	0.4	0.1	0.8	A
2	PLANK	PLNK-PARM2	CMMLT	0.00001	0.000001	0.1	M
2	PLANK	PLNK-PARM2	CMMN	0.025	0.0025	0.25	M
2	PLANK	PLNK-PARM2	CMMP	0.005	0.0005	0.05	M
2	PLANK	PLNK-PARM2	TALGRH	150	100	200	A
2	PLANK	PLNK-PARM2	TALGRL	-100	0	-118	A
2	PLANK	PLNK-PARM2	TALGRM	70	50	90	A
2	PLANK	PLNK-PARM3	ALR20	0.003	0.0003	0.03	M
2	PLANK	PLNK-PARM3	ALDH	0.002	0.0002	0.02	M
2	PLANK	PLNK-PARM3	ALDL	0.002	0.0002	0.02	M
2	PLANK	PLNK-PARM3	OXALD	0.03	0.003	0.3	M
2	PLANK	PLNK-PARM3	NALDH	0.001	0.0001	0.01	M
2	PLANK	PLNK-PARM3	PALDH	0.0001	0.00001	0.001	M
2	PLANK	PHYTO-PARM	SEED	0.2	0.02	2	M
2	PLANK	PHYTO-PARM	MXSTAY	0.2	0.02	2	M
2	PLANK	PHYTO-PARM	OREF	1000	100	10000	M
2	PLANK	PHYTO-PARM	CLALDH	50	5	500	M
2	PLANK	PHYTO-PARM	PHYSET	0.02	0.001	0.3	M
2	PLANK	PHYTO-PARM	REFSET	0.1	0.001	1	M
2	PLANK	BENAL-PARM	MBAL	100000	1000	1000000	M

2	PLANK	BENAL-PARM	CFBALR	0.4	0.1	0.8	A
2	PLANK	BENAL-PARM	CFBALG	0.4	0.1	0.8	A